CHALLENGES ON THE PATH TO REGULAR ONLINE SHOPPING: E-GROCERY SECTOR - UK

Pedro Ferreira
University of Reading / Universidade Católica – Porto / Sponsor: FCT
University of Reading, School of Business, P.O. Box 218, RG6 6AA, Reading, UK
Ler00mpf@reading.ac.uk

ABSTRACT

This paper attempts to uncover significant economic and non-economic demand-side variables, which are driving and hindering B2C (Business to Consumer) e-commerce learning. It investigates the perceptions of individual customers on the path towards a specific regular type of online-buying: E-Grocery shopping (EGS).
The analysis and result display is based on an e-customer learning framework, consisting of clear crucial steps arranged in a tree decision format which illustrate decisions faced by customers as they evolve from non-Internet user to regular E-grocery shopper. This framework was applied via mail survey to a sample of 2036 households in England.
Results (through regression and mean testing) are pointing at two critical barriers on the path to E-grocery:

- One network specific: National Digital Divide creating an important cost for potential users.
- One sector specific: E-grocery is a sector plagued by non-friendly sites together with deficient logistics.

With this cost structure, the niche market for e-grocers seems to be a reality, confirming Pfeffers [3] view. Income proves to be a key variable behind e-shopping learning, generating a very high and cumulative premium tag on e-grocery. Also interesting, is the fact that grocery shopping still preserves its feminine connotation online.

KEYWORDS

E-Grocery, E-commerce, Income, Learning, Niche market.

1. THEORETICAL FRAMEWORK

Buying groceries online is quite complicated for a potential customer. It is a type of B2C shopping that involves a high degree of familiarity with surfing tools and potential faults. Setting up a regular shopping list and customer account in an e-grocers site is not an easy task and normally will take more than 30 minutes. Also, delivery takes time to be arranged in a proficient way and may be prone to delays.

Any customer that wants to buy groceries online, must first become familiar with the platform where the transaction takes place, naturally proceeding from Internet user to casual buyer and finally to regular shopper.

Going through these stages puts potential customers into trouble and takes time. Whenever they proceed into a new stage they are faced with different learning challenges which present both advantages and disadvantages. Successfully tackling each stage involves having enough drive or motivation to compensate for learning costs and potential faults.

Yrjöla, Tanskanen and Holmström [6], tried to uncover the reasons behind the slowly uptake and even failure of e-grocers. Looking at the demand side, these authors consider that a significant part of the problem is probably determined by demand itself, because, even though customers are looking for convenience, "it takes some time to learn new buying routines before the time-saving element is fully appreciated.", page 6.

Regardless of their business model, this situation may restrict e-grocers to a niche market, keeping them from thriving (especially those with fully automated models hungry for mass-markets). The fast sequence of Internet grocery start-up failures in the US (Webvan, Streamline, Kozmo, Homeruns, among others) during 2000-2001, alongside with the still non-profitability of most operators elsewhere, including UK, provide evidence of how tough ambitious strategies can be.

At the actual state of play, operators targeting the E-grocery mass-market must be willing to internalise specific and significant demand costs that are restricting potential for growth in this sector.

In order to assess the magnitude and extent of costs faced and perceived by the demand side in the UK, research will focus on observing how UK households respond to barriers and motivators as they move from an offline to a regular online buying status. Attention will be paid to important variables believed to be conditioning household behaviour. The following features: gender, type of occupation, age, number of people and children in household, number of cars owned, household income and population density in area are expected to play a significant role decision-wise. Servon [2], page 1, on a study mainly focusing upon the US reality, believes that: "In virtually all countries, Internet users tend to be young, urban, male, and relatively well educated and wealthy". This profile also seems to apply to some extent to the UK, with Internet users being typically male and young, according to the UK National Statistics [5], page 2.

From the above, income will be closely monitored in order to determine to what extent E-grocery shopping is a premium sector in a premium market. It will also be interesting to distinguish between costs, which are network specific (Internet), activity specific (Buying Online) and sector specific (Online Grocery).

2. SAMPLING

To start shedding some light upon the above issues, a demand mail survey was conducted in early March 2003, targeting 2036 households in England. The database used was BT's phone disc [4], containing all UK fixed phone numbers and addresses. The households were sampled via a weighing procedure, which allocated observations to postcodes within Local Authorities (LAs), according to their total number of households. Final selection within postcodes was random based on the first two letters of surnames from the database. This set of procedures tried to create a sample that could be as representative as possible, by: being random, covering all postcodes within LAs and capturing the presence of ethnic minorities where available.

Local Authorities (three from London (urban) and two from the West Midlands (rural)) covered in the survey were selected based on the diversity they presented in terms of income and density.

A specific bias favouring high-income areas is deliberately present to avoid a low proportion of e-grocery shoppers (also Internet users) that could severely reduce significance of any analysis at the sector level. This bias can be seen in a final figure of 70% for Internet access in the sample compared with 62% (in a broader view) claimed in February 2003 by the UK National Statistics [5] page 3. Other database specific biases are also in place; biases like preponderance of male in the sample and reduced percentage of youngsters, i.e., 16 to 25 years old, are due to the database itself and to the unit under exam: the household. Fixed telephone lines have traditionally been registered in the man's name and are requested by mature house owners, not trendy youngsters with no jobs, no money, studying or living with parents and using mobile phones.

Biases may distort the significance of figures, so to reduce their effect, emphasis was primarily placed upon the rationale behind decisions rather than upon the bare absolute and relative numbers at each stage.

3. RESULTS

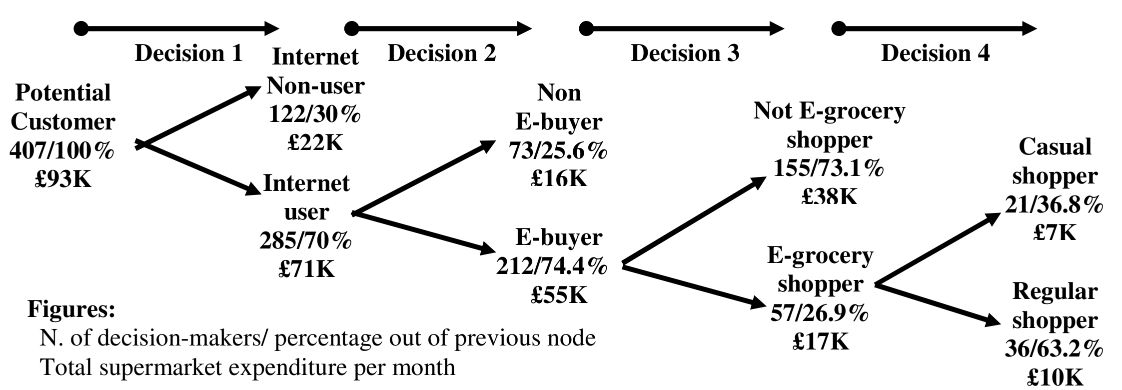

Figure 1. Sample decision tree on the way to Regular E-grocery shopping.

Figure 1 displays the tree structure guiding questionnaire design and presents the first results.

The categories displayed here were devised in the following manner: User (someone that uses the Internet in a regular basis, at least once a month), Buyer on B2C (someone that actually spent money online at anytime in the past), E-grocery shopper (someone that used an e-grocery service, at least once) and regular e-grocery shopper (someone that buys e-groceries monthly or more than 5 times overall).

By following the decision tree from left to right, it is quite clear that the costs of learning start to build up to such an extent that only a few can make it into the regular online grocery shopping stage. This implies an increasing difficulty in learning that accumulates, as Internet activities become more complicated and risky.

On a decisional basis, it seems like decisions 1 and 3 are the two most significant causes of drop-out, however, they tend to affect learning in different ways. Decision 1 is a critical and global decision for any user, determining whether or not he or she is included or excluded from any learning process. Exclusion means discrimination in accessing extensive info sources and opportunities only available electronically.

Decision 3, even though being by far the most costly one causing a drop out ratio of approximately 7 in every group of 10 individuals, it does not present such an immediate overall impact upon general E-commerce learning, as it is more of a sector choice. Probably, this drop-out ratio only emphasises the complex nature of this activity and the low level of trust inspired by this still novel sector.

Decision 2 represents a smoother step on the decision tree, probably displaying the fact that around 64% of Internet users in the sample perceive convenience (time saving plus easiness) in online buying.

In a glance, all the above costs may even be more significant due to underestimation of droppers (bias-2.).

4. LOOKING UNDERNEATH THE TREE: MODELLING

To learn more about each decision and about the framework itself, an attempt was made to do an econometric analysis by regressing each one of the decisions against specific explanatory variables proposed earlier (1.)

The regression model chosen was the Sequential Logit Model. Logit (All models include constant: $Logit(Y) = B_1 + B_2*X_1 + \ldots + B_n*X_n$), because the dependent variables are qualitative binomials and tend to be concentrated in tails (70/30 splits). Sequential, because the decision tree must be respected. So when there is a movement down the tree, samples must be restricted to sub-samples for further regressions.

The final modelling was based on a general to specific procedure trying to arrive to a best-fit specification. Starting with all explanatory variables in the model, exclusion of variables was applied in order to correct for eventual auto-correlation, lack of significance and others. Testing for this modelling was realised via Limdep 7 and SPSS 11 softwares leading to similar results, whether, stepwise forward or backwards. Confidence level for acceptance and variable exclusion was 95%.

Table 1. Initial regressions along with best fit models for each decision
* Gender: female=1 and male=2 ** Occupation: Full-Time=1, Part-time=2, Retired=3 and Unemployed=4

95% Confidence Level	**Decision 1** Internet usage		**Decision 2** Buying Online		**Decision 3** Buy Online Groc		**Decision 4** Buy Regular Groc	
Valid N – Sample	316		230		178		45	
Explanatory Variables	**Coeff.**	**Sign.**	**Coeff.**	**Sign.**	**Coeff.**	**Sign.**	**Coeff.**	**Sign.**
Gender*	.312	No	-.196	No	-.56	No	.236	No
Occupation**	-.277	No	.524	No	.679	No	.396	No
Age	**-1.057**	**Yes**	**-.624**	**Yes**	-.393	No	.227	No
Number of People	-.017	No	-.092	No	-.184	No	.004	No
Number of Children	-.105	No	-.084	No	.461	No	.344	No
Number of Cars	**.743**	**Yes**	**.581**	**Yes**	.168	No	.21	No
Income	**.425**	**Yes**	**.326**	**Yes**	**.563**	**Yes**	-.176	No
Density	.279	No	.045	No	-.145	No	.926	No

Best fit models:

Decision 1: Logit(Use) = 3.151-1.162*Age+0.664*Cars+0.492*Income

Decision 2: Logit(Buy) = 0.041+.319*Income

Decision 3: Logit(Groc) = 1.613 –0.791*Gender+0.379*Children+0.403*Income

Decision 4: No significance allowing for best fit – small sample and insufficient variance.

Changes in significance between full regressions and best-fit models in table 1 are essentially due to corrections of correlation displayed between explanatory variables.

As expected, income proved to be the most significant variable driving the general e-commerce learning process. Income is significant in all decisions. Potential learners with higher income levels can move smoothly down the decision tree, value more convenience and easily face two kinds of costs: costs with learning, i.e., hardware, software, Internet access, knowledge, and costs with risky choices, i.e., new online challenges like buying. This result was also obtained by Lunn and Suman[1], page 567, as "… individuals with higher incomes tend to purchase more frequently and spend more money on the Internet"

The premium character of the network and sector seems to be confirmed by this result. At the sector level, the premium barrier is further reinforced by the accumulation of income restrictions from decisions 1 and 2.

The threshold for E-grocery, however, is not only dependent on income and is further magnified by other significant variables. At the platform level, by age, as users tend to be mostly young adults. At the sector level, by gender and number of children in household, as e-grocery shopping still preserves a feminine and family-oriented connotation going against the dominant online gender: male. On the whole, the combination and accumulation of these particular features leaves a small market niche for actual market players.

In the end, the profile identified for a regular E-Grocery shopper is the following, see table 2:

Table 2. Profile for regular E-grocery shoppers, determined via mean analysis (sample T-test at 95% C Level)

	Profiling Variables							
	Gender	**Occupation**	**Age**	**People**	**Children**	**Cars**	**Income**	**Density**
Reg Groc	Female	Part-Time	26-44	4	2	1-2	100K +	V. Hi

5. CONCLUSIONS

At the moment, learning on the way to regular electronic shopping is a premium cumulative activity.

On a decisional basis, the greatest learning cost is sector based, as imperfections seem to be taking their toll.

Platform costs associated with Internet usage should not also be forgotten as they portray a situation of overall virtual market constrained by national digital divide, which is important in this sample.

With the above in mind, if E-grocers want to target customers online, they can only expect to have a niche market. So far, authors like Ken Peffers[3] seem to be getting it right. He believes that presently grocery items do not possess the required features to generate a successful online market. They are seen as too bulky relative to their value. To be sold online, these goods, due to their high non-codified info content and physical and perishable nature, involve high costs with online specification and home delivery. Therefore, they deliver low profit margins and require high levels of investment in delivery systems that are still away from being customer friendly on a large scale.

6. REFERENCES

Book

[1] Lunn, R. and Suman, M., 2002, *Experience and Trust in Online Shopping*, from *The Internet in Everyday Life*, edited by Wellman, B. and Haythornthwaite, C., Blackwell Publishing.

[2] Servon, L., 2002, *Bridging the Digital Divide: Technology, Community, and Public Policy*, Blackwell Publishing.

Journal

[3] Peffers, K., 2001, *The Future of Electronic Commerce: A Shift from the EC Channel to Strategic Electronic Commerce*, Journal of Information Technology Theory and Application, 3:4, 7-16.

Conference paper, Internet and other

[4] British Telecom, 2002, *BT Phone Disc – 3.2 Standard Version*, www.bt.com.

[5] UK National Statistics, April 2003, *Internet Access: Households and Individuals*, www.statistics.gov.uk/releases.

[6] Yrjölä, H., et al, 2000, *The Way to Profitable Internet Grocery Retailing – 6 Lessons Learned*, http://www.tuta.hut.fi/ecomlog/publications.